

# On Optimal Causal Coding of Partially Observed Markov Sources in Single and Multi-Terminal Settings

Serdar Yüksel<sup>1</sup>

**Abstract**—The optimal causal (zero-delay) coding of a partially observed Markov process is studied, where the cost to be minimized is a bounded, non-negative, additive, measurable single-letter function of the source and the receiver output. A structural result is obtained extending Witsenhausen's and Walrand-Varaiya's structural results on optimal causal coders to more general state spaces and to a partially observed setting. The decentralized (multi-terminal) setup is also considered. For the case where the source is an i.i.d. process, it is shown that an optimal solution to the decentralized causal coding of correlated observations problem is memoryless. For Markov sources, a counterexample to a natural separation conjecture is presented.

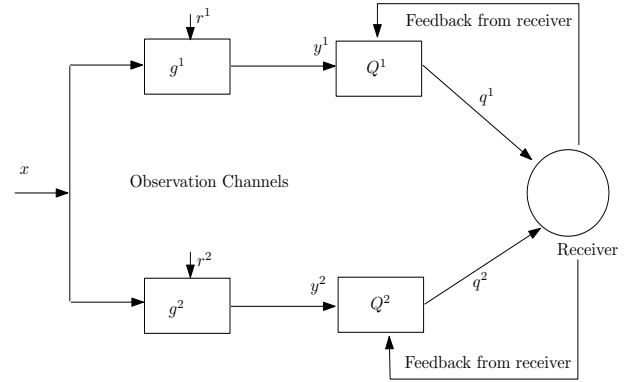


Fig. 1: Partially observed source under a decentralized structure.

## I. INTRODUCTION

This paper considers optimal causal encoding/quantization of partially observed Markov processes. We begin with providing a description of the system model. We consider a partially observed Markov process, defined on a probability space  $(\Omega, \mathcal{F}, P)$  and described by the following discrete-time equations for  $t \geq 0$ :

$$x_{t+1} = f(x_t, w_t), \quad (1)$$

$$y_t^i = g^i(x_t, r_t^i), \quad (2)$$

for (Borel) measurable functions  $f, g^i, i = 1, 2$ , with  $\{w_t, r_t^i, i = 1, 2\}$  i.i.d., mutually independent noise processes and  $x_0$  a random variable with probability measure  $\nu_0$ . Here, we let  $x_t \in \mathbb{X}$ , and  $y_t^i \in \mathbb{Y}^i$ , where  $\mathbb{X}, \mathbb{Y}^i$  are complete, separable, metric spaces (Polish spaces), and thus, include countable spaces or  $\mathbb{R}^n, n \in \mathbb{N}_+$ .

Let an encoder, Encoder  $i$ , be located at one end of an observation channel characterized by (2). The encoders transmit their information to a receiver (see Figure 1), over a discrete noiseless channel with finite capacity; that is, they quantize their information. The information at the encoders may also contain feedback from the receiver, which we clarify in the following.

Let us first define a quantizer.

<sup>1</sup>Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Research supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Email: yuksel@mast.queensu.ca. This work was presented in part at the at the IEEE International Symposium on Information Theory, Austin, Texas, in 2010. Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

**Definition 1.1:** Let  $\mathcal{M} = \{1, 2, \dots, M\}$  with  $M = |\mathcal{M}|$ . Let  $\mathbb{A}$  be a topological space. A quantizer  $Q(\mathbb{A}; \mathcal{M})$  is a Borel measurable map from  $\mathbb{A}$  to  $\mathcal{M}$ .  $\diamond$

When the spaces  $\mathbb{A}$  and  $\mathcal{M}$  are clear from context, we will drop the notation and denote the quantizer simply by  $Q$ .

We refer by a **Composite Quantization Policy**  $\Pi^{comp, i}$  of Encoder  $i$ , a sequence of functions  $\{Q_t^{comp, i}, t \geq 0\}$  which are causal such that the quantization output at time  $t$ ,  $q_t^i$ , under  $\Pi^{comp, i}$  is generated by a causally measurable function of its local information, that is, a mapping measurable on the sigma-algebra generated by

$$I_t^i = \{y_{[0, t]}^i, q_{[0, t-1]}^1, q_{[0, t-1]}^2\}, \quad t \geq 1,$$

and  $I_0^i = \{y_0^i\}$ , to a finite set  $\mathcal{M}_t^i$ , the quantization output alphabet at time  $t$  given by

$$\mathcal{M}_t^i := \{1, 2, \dots, |\mathcal{M}_t^i|\},$$

for  $0 \leq t \leq T - 1$  and  $i = 1, 2$ . Here  $\{\mathcal{M}_t^i\}$  are fixed in advance and do not depend on the realizations of the random variables. Here, we have the notation for  $t \geq 1$ :

$$y_{[0, t-1]}^i = \{y_s^i, 0 \leq s \leq t - 1\}.$$

Let

$$\mathbb{I}_t^i = \left( \prod_{s=0}^{t-1} \mathcal{M}_s^1 \times \mathcal{M}_s^2 \right) \times (\mathbb{Y}^i)^{t+1}, \quad t \geq 1, \quad \mathbb{I}_0^i = \mathbb{Y}^i,$$

be information spaces such that for all  $t \geq 0$ ,  $I_t^i \in \mathbb{I}_t^i$ . Thus,

$$Q_t^{comp, i} : \mathbb{I}_t^i \rightarrow \mathcal{M}_t^i.$$

We may express, equivalently, the policy  $\Pi^{comp,i}$  as a composition of a **Quantization Policy**  $\Pi^i$  and a **Quantizer**. A quantization policy of encoder  $i$ ,  $\mathcal{T}^i$ , is a sequence of functions  $\{T_t^i\}$ , such that for each  $t \geq 0$ ,  $T_t^i$  is a mapping from the information space  $\mathbb{I}_t^i$  to a space of quantizers  $\mathbb{Q}_t^i$ . A quantizer, subsequently is used to generate the quantizer output. That is for every  $t$  and  $i$ ,  $T_t^i(I_t^i) \in \mathbb{Q}_t^i$  and for every  $I_t^i \in \mathbb{I}_t^i$ , we will adopt a following representation

$$Q_t^{comp,i}(I_t^i) = (T_t^i(I_t^i))(I_t^i), \quad (3)$$

mapping the information space to  $\mathcal{M}_t^i$  in its most general form. We note that even though there may seem to be duplicated information in (3) (since a map is used to pick a quantizer, and the quantizer maps the available information to outputs) we will eliminate any informational redundancy: A quantizer action will be generated based on the common information at the encoder and the receiver, and the quantizer will map the relevant private information at the encoder to the quantization output. Such a separation in the design will also allow us to use the machinery of Markov Decision Processes to obtain a structural method to design optimal quantizers, to be clarified further, without any loss in optimality.

That is, let the information at the receiver at time  $t \geq 0$  be  $I_t^r = \{q_{[0,t-1]}^1, q_{[0,t-1]}^2\}$ . The common information, under feedback information, in the encoders and the receiver is the set  $I_t^r \in \left(\prod_{s=0}^{t-1} \mathcal{M}_s^1 \times \mathcal{M}_s^2\right)$ . Thus, we can express the composite quantization policy as:

$$Q_t^{comp,i}(I_t^i) = (T_t^i(I_t^r))(I_t^i \setminus I_t^r), \quad (4)$$

mapping the information space to  $\mathcal{M}_t^i$ . We note that, any composite quantization policy  $Q_t^{comp,i}$  can be expressed in the form above; that is there is no loss in the space of possible such policies, since for any  $Q_t^{comp,i}$ , one could define

$$T_t^i(I_t^r)(\cdot) := Q_t^{comp,i}(I_t^r, \cdot).$$

Thus, we let  $\text{DM}^i$  have policy  $\mathcal{T}^i$  and under this policy generate quantizer actions  $\{Q_t^i, t \geq 0\}$ ,  $Q_t^i \in \mathbb{Q}_t^i$  ( $Q_t^i$  is the quantizer used at time  $t$ ). Under action  $Q_t^i$ , and given the local information, the encoder generates  $q_t^i$ , as the *quantization output* at time  $t$ .

The receiver, upon receiving the information from the encoders, generates its decision at time  $t$ , also causally: An admissible causal receiver policy is a sequence of measurable functions  $\gamma = \{\gamma_t\}$  such that

$$\gamma_t : \prod_{s=0}^t \left(\mathcal{M}_s^1 \times \mathcal{M}_s^2\right) \rightarrow \mathbb{U}, \quad t \geq 0$$

where  $\mathbb{U}$  denotes the decision space.

For a general vector  $a$ , let  $\mathbf{a}$  denote  $\{a^1, a^2\}$  and let  $\Pi = \{\Pi^1, \Pi^2\}$  denote the ensemble of policies and  $\mathbf{Q}_t = \{Q_t^1, Q_t^2\}$ . Hence,  $\mathbf{q}_{[0,t]}$  denotes  $\{q_{[0,t]}^1, q_{[0,t]}^2\}$ .

With the above formulation, the objective of the decision makers is the following minimization problem:

$$\inf_{\Pi^{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi^{comp}, \gamma} \left[ \sum_{t=0}^{T-1} c(x_t, v_t) \right],$$

over all policies  $\Pi^{comp}$ ,  $\gamma$  with the random initial condition  $x_0$  having probability measure  $\nu_0$ . Here  $c(\cdot, \cdot)$ , is a non-negative, bounded, measurable function and  $v_t = \gamma_t(\mathbf{q}_{[0,t]})$  for  $t \geq 0$ .

We also assume that the encoders and the receiver know the apriori distribution  $\nu_0$ .

Before concluding this section, it may be worth emphasizing the operational nature of causality; as different approaches have been adopted in the literature. The encoders at any given time can only use their local information to generate the quantization outputs. The receiver, at any given time, can only use its local information to generate its decision/estimate. These happen with zero delay, that is if there is a common-clock at the encoders and the receiver; the receiver at time  $t$  needs to make its decision before the realizations  $x_{t+1}, y_{t+1}^1, y_{t+1}^2$  have taken place. This corresponds to the zero-delay coding schemes of, for example, Witsenhausen and Linder-Lugosi in [52], [29]; but is different from the setup of Neuhoff and Gilbert [40], which allows long delays at the decoder. Our motivation for such zero-delay schemes comes mainly from applications in networked control systems, when sensors need to transmit information to controllers who need to act on a system; such systems cannot tolerate long delays, in particular when the source is open loop unstable and disturbance exists in the evolution of the source.

#### A. Relevant literature

Some related studies of the above setup include optimal control with multiple sensors and sequential decentralized hypothesis testing problems and multi-access communications with feedback [3]. Related papers on real-time coding include the following: [40] established that the optimal optimal causal encoder minimizing the data rate subject to a distortion for an i.i.d sequence is memoryless. If the source is  $k$ th-order Markov, then the optimal causal fixed-rate coder minimizing any measurable distortion uses only the last  $k$  source symbols, together with the current state at the receiver's memory [52]. Reference [46] considered the optimal causal coding problem of finite-state Markov sources over noisy channels with noiseless feedback. [43] and [37] considered optimal causal coding of Markov sources over noisy channels without feedback. [36] considered the optimal causal coding over a noisy channel with noisy feedback. Reference [30] considered the causal coding of stationary sources under a high-rate assumption.

Our paper is particularly related to the following two efforts in the literature: Borkar, Mitter and Tatikonda [11] studies a related problem of coding of a partially observed Markov source, however, the construction for the encoders is restricted to take a particular form which uses the information at the decoder and the most recent observation at the encoder (not including the observation history). As another point of relevance with our paper, [11] regarded the actions as the *quantizer functions*, which we will discuss further. In contrast, the only restriction we have in this paper is causality, in the zero-delay sense. On the other hand, we do not claim the existence results that the authors in [11] are making. Another work in the literature which is related to ours is by Nayyar and Teneketzis [38], considering a multi-terminal setup. [38] considers decentralized coding of correlated sources when the encoders observe

conditionally independent messages given a finitely valued random variable and obtain separation results for the optimal encoders. The paper also considers noisy channels. In our setup, there does not exist a finitely valued random variable which makes the observations at the encoders conditionally independent.

References [47] and [31] consider optimal causal variable-rate coding under side information and [55] considers optimal variable-rate causal coding under distortion constraints. The studies in [31] and [55] are in the context of real-time, zero-delay settings; whereas [47] considers causality in the sense of Neuhoﬀ and Gilbert [40] as discussed in the previous section.

We will also obtain structural results for optimal decentralized coding of i.i.d. sources. There are algorithmic results available in the literature when the encoders satisfy the optimal structure obtained in the paper, important resources in this direction include [24], [21] and [48].

A parallel line of consideration which has a rate-distortion theoretic nature is on *sequential-rate distortion* proposed in [42] and the *feedforward* setup, which has been investigated in [45] and [17].

Our work is also related to Witsenhausen's indirect rate distortion problem [51] (see also [15]). We will observe that, the separation argument through the *disconnection principle* of [51] applies to our setting in a dynamic context. Further related papers include [6], [28].

In our paper, we also use ideas from team decision theory, see [50], [54], [34], [35] and [13] for related discussions and applications.

## B. Contributions of the paper

- The optimal causal coding of a partially observed Markov source is considered. For the single terminal case, a structural result is obtained extending Witsenhausen's and Walrand and Varaiya's structural results on optimal causal (zero-delay) coders to a partially observed setting and to sources which take values in a Polish space. We show that a separation result of a form involving the decoder's belief on the encoder's belief on the state is optimal.
- The decentralized (multi-terminal) setup is also considered. For the case where the source is an i.i.d. process, it is shown that the optimal decentralized causal coding of correlated observations problem admits a solution which is memoryless. For Markov sources, a counterexample to a natural separation conjecture is presented. The decentralized control concept of *signaling* is interpreted in the context of decentralized coding.
- The results are applied to a Linear-Quadratic-Gaussian (LQG) estimation/optimization problem. The results above induce an optimality result for *separation of estimation and quantization*, where the estimation is obtained with a Kalman Filter and the filter output is quantized.

We now summarize the rest of the paper. In section II, we present our results on optimal coding of a partially observed Markov process when there is only one encoder. Section III discusses the decentralized setting for a multi-encoder setup and presents a counterexample for a separation conjecture and

provides a separation result when the source is memoryless. The paper ends with the concluding remarks of section V, following an application example on linear, Gaussian systems in Section IV. The proofs of the results are presented in the Appendix.

## II. SINGLE-TERMINAL CASE: OPTIMAL CAUSAL CODING OF A PARTIALLY OBSERVED MARKOV SOURCE

### A. Revisiting the single-terminal, fully observed case

Let us revisit the single-encoder, fully observed case: In this setup,  $y_t = x_t$  for all  $t \geq 0$ . There are two related approaches in the literature as presented explicitly by Teneketzis in [43]; one adopted by Witsenhausen [52] and one by Walrand and Varaiya [46]. Reference [43] extended the setups to the more general context of non-feedback communication.

**Theorem 2.1:** [Witsenhausen [52]] Any (*causal*) *composite quantization policy* can be replaced, without any loss in performance, by one which only uses  $x_t$  and  $q_{[0,t-1]}$  at time  $t \geq 1$ .  $\diamond$

Walrand and Varaiya considered sources living in a finite set, and obtained the following:

**Theorem 2.2:** [Walrand-Varaiya [46]] Any optimal (*causal*) *composite quantization policy* can be replaced, without any loss in performance, by one which only uses the conditional probability measure  $P(x_t|q_{[0,t-1]})$ , the state  $x_t$ , and the time information  $t$ , at time  $t \geq 1$ .  $\diamond$

The difference between the structural results above is the following: In the setup suggested by Theorem 2.1, the encoder's memory space is not fixed and keeps expanding as the decision horizon in the optimization,  $T - 1$ , increases. In Theorem 2.2, the memory space of an optimal encoder is fixed. In general, the space of probability measures is a very large one; however, it may be the case that different quantization outputs may lead to the same conditional probability measure on the state process, leading to a reduction in the required memory. Furthermore, Theorem 2.2 allows one to apply the theory of Markov Decision Processes, for infinite horizon problems. We note that [11] applied such a machinery to obtain existence results for optimal causal coding of partially observed Markov processes

### B. Optimal causal coding of a partially observed Markov source

Consider the setup earlier in (2) with a single encoder. Thus, the system considered is a discrete-time scalar system described by

$$x_{t+1} = f(x_t, w_t), \quad y_t = g(x_t, r_t), \quad (5)$$

where  $x_t$  is the state at time  $t$ , and  $\{w_t, r_t\}$  is a sequence of zero-mean, mutually independent, identically distributed (i.i.d.) random variables with finite second moments. Let the quantizer, as described earlier, map its information to a finite set  $\mathcal{M}_t$ . At any given time, the receiver generates a quantity  $v_t$  as a function of its received information, that is as a function of  $\{q_0, q_1, \dots, q_t\}$ . The goal is to minimize  $\sum_{t=0}^{T-1} E[c(x_t, v_t)]$ , subject to constraints on the number of quantizer bins in  $\mathcal{M}_t$ , and the causality restriction in encoding and decoding.

Let for a Polish space  $\mathbb{S}$ ,  $\mathcal{P}(\mathbb{S})$  be the space of probability measures on  $\mathcal{B}(\mathbb{S})$ , the Borel  $\sigma$ -field on  $\mathbb{S}$  (generated by open sets in  $\mathbb{S}$ ). At this point pause to provide a brief discussion on the space  $\mathcal{P}(\mathbb{S})$ .

Let  $\Gamma(\mathbb{S})$  be the set of all Borel measurable and bounded functions from  $\mathbb{S}$  to  $\mathbb{R}$ . We first wish to find a topology on  $\mathcal{P}(\mathbb{S})$ , under which functions of the form:

$$\Theta := \left\{ \int_{x \in \mathbb{S}} \pi(dx) f(x), \quad f \in \Gamma(\mathbb{S}) \right\}$$

are measurable on  $\mathcal{P}(\mathbb{S})$ . We will need this to construct the structure of optimal quantizers later in this section.

Let  $\{\mu_n, n \in \mathbb{N}\}$  be a sequence in  $\mathcal{P}(\mathbb{S})$ . Recall that  $\{\mu_n\}$  is said to converge to  $\mu \in \mathcal{P}(\mathbb{S})$  *weakly* if

$$\int_{\mathbb{S}} c(x) \mu_n(dx) \rightarrow \int_{\mathbb{S}} c(x) \mu(dx)$$

for every continuous and bounded  $c : \mathbb{S} \rightarrow \mathbb{R}$ . The sequence  $\{\mu_n\}$  is said to converge to  $\mu \in \mathcal{P}(\mathbb{S})$  *setwise* if

$$\int_{\mathbb{S}} c(x) \mu_n(dx) \rightarrow \int_{\mathbb{S}} c(x) \mu(dx)$$

for every measurable and bounded  $c : \mathbb{S} \rightarrow \mathbb{R}$ . For two probability measures  $\mu, \nu \in \mathcal{P}(\mathbb{S})$ , the *total variation* metric is given by

$$\begin{aligned} \|\mu - \nu\|_{TV} &:= 2 \sup_{B \in \mathcal{B}(\mathbb{S})} |\mu(B) - \nu(B)| \\ &= \sup_{f: \|f\|_{\infty} \leq 1} \left| \int f(x) \mu(dx) - \int f(x) \nu(dx) \right|, \end{aligned}$$

where the infimum is over all measurable real  $f$  such that  $\|f\|_{\infty} = \sup_{x \in \mathbb{S}} |f(x)| \leq 1$ . A sequence  $\{\mu_n\}$  is said to converge to  $\mu \in \mathcal{P}(\mathbb{S})$  in total variation if  $\|\mu_n - \mu\|_{TV} \rightarrow 0$ .

These three convergence notions are in increasing order of strength: convergence in total variation implies setwise convergence, which in turn implies weak convergence. Total variation is a very strong notion for convergence. Furthermore, the space of probability measures under total variation metric is not separable. Setwise convergence also induces an inconvenient topology on the space of probability measures, particularly because this topology is not metrizable ([20, p. 59]). However, the space of probability measures on a complete, separable, metric (Polish) space endowed with the topology of weak convergence is itself a complete, separable, metric space [4]. The Prohorov metric, for example, can be used to metrize this space, among other metrics. This topology has found many applications in information theory and stochastic control. For these reasons, one would like to work with weak convergence.

By the above definitions, it is evident that both setwise convergence and total variation are sufficient for measurability of the function class  $\Theta$ , since under these topologies  $\int \pi(dx) f(x)$  is (sequentially) continuous on  $\mathcal{P}(\mathbb{S})$  for every  $f \in \Gamma(\mathbb{S})$ . However, as we state in the following, weak convergence is also sufficient (see Theorem 15.13 in [4] or p. 215 in [8]).

**Theorem 2.3:** Let  $\mathbb{S}$  be a Polish space and let  $M(\mathbb{S})$  be the set of all measurable and bounded functions  $f : \mathbb{S} \rightarrow \mathbb{R}$  under which

$$\int \pi(dx) f(x)$$

defines a measurable function on  $\mathcal{P}(\mathbb{S})$  under the weak convergence topology. Then,  $M(\mathbb{S})$  is the same as  $\Gamma(\mathbb{S})$  of all bounded and measurable functions.  $\diamond$

Hence,  $\mathcal{P}(\mathbb{S})$  will denote the space of probability measures on  $\mathbb{S}$  under weak convergence. Now, define  $\pi_t \in \mathcal{P}(\mathbb{X})$  to be the regular conditional probability measure given by

$$\pi_t(A) = P(x_t \in A | y_{[0,t]}), \quad A \in \mathcal{B}(\mathbb{X}).$$

The existence of this regular conditional probability measure for every realization  $y_{[0,t]}$  follows from the fact that both the state process and the observation process are Polish. It is known that the process  $\{\pi_t\}$  evolves according to a non-linear filtering equation (see (14)), and is itself a Markov process (see [10], [49], [11]).

Let us also define  $\Xi_t \in \mathcal{P}(\mathcal{P}(\mathbb{X}))$  as the regular conditional measure

$$\Xi_t(A) = P(\pi_t \in A | q_{[0,t-1]}), \quad A \in \mathcal{B}(\mathcal{P}(\mathbb{X})).$$

The following are the main results of this section:

**Theorem 2.4:** Any (causal) composite quantization policy can be replaced, without any loss in performance, by one which only uses  $\{\pi_t, q_{[0,t-1]}\}$  as a sufficient statistic for  $t \geq 1$ . This can be expressed as a quantization policy which only uses  $q_{[0,t-1]}$  to generate a quantizer, where the quantizer uses  $\pi_t$  to generate the quantization output at time  $t$ .  $\diamond$

**Theorem 2.5:** Any optimal (causal) composite quantization policy can be replaced, without any loss in performance, by one which only uses  $\{\Xi_t, \pi_t, t\}$  for  $t \geq 1$ . This can be expressed as an optimal quantization policy which only uses  $\{\Xi_t, t\}$  to generate an optimal quantizer, where the quantizer uses  $\pi_t$  to generate the quantization output at time  $t$ .  $\diamond$

The proofs of the results above are presented in the Appendix. We present two remarks in the following:

**Remark 2.1:** Our results above are not surprising. In fact, once one recognizes the fact that  $\{\pi_t\}$  forms a Markov source, and the cost function can be expressed as a function  $\tilde{c}(\pi, v)$ , for some function  $\tilde{c} : \mathcal{P}(\mathbb{X}) \times \mathbb{U} \rightarrow \mathbb{R}$ , one could almost directly apply Witsenhausen's [52] as well as Walrand and Varaiya's [46] results to recover the structural results above (except the fact that Walrand and Varaiya consider sources living in a finite alphabet). The proofs in the Appendix are presented for completeness and to address the technical intricacies.  $\diamond$

**Remark 2.2:** Having the actions as the quantizers, and not the quantizer outputs, allows one to define a Markov Decision Problem with well-defined cost functions and state and action spaces. By the proof of Theorem 2.5, we will observe that  $(\Xi, Q_t)$  forms a Markov Chain, which is a key observation: Thus, the action space can be constructed as some topological space of quantizers acting on  $\mathcal{P}(\mathbb{X})$ . Borkar, Mitter and Tatikonda [11] adopted this view while formulating an MDP optimization problem, where the quantizer acts on  $\mathbb{Y}$  (As mentioned earlier, our separation result is different from [11] due to the structure imposed on the quantizers in [11]). See also [56] and [58] for a topology on quantizers.  $\diamond$

### C. Extensions to finite delay decoding and higher order Markov sources

The results presented are also generalizable to settings where, a) the source is Markov of order  $m > 0$ , b) a finite delay  $d$  is allowed at the decoder, and c) the observation process depends also on past source outputs in a sense described in (6) below. For these cases, we consider the following generalization of the source by expanding the state space.

Suppose that the partially observed source is such that, the source is Markov of order  $m$ , or there is a finite delay  $d > 0$  which is allowed at the decoder. In this case, we can augment the source to obtain  $z_t = \{x_{[t-\max(d+1,m)+1,t]}\}$ . Note that  $\{z_t\}$  is Markov. We can thus consider the following representation:

$$z_{t+1} = f(z_t, w_t), \quad y_t = g(z_t, r_t), \quad (6)$$

where  $z_t = \{x_{[t-\max(d+1,m)+1,t]}\} \in \mathbb{X}^{\max(d+1,m)}$ , and  $r_t, w_t$  are mutually independent, i.i.d. processes.

Any per-stage cost function of the form  $c(x_t, v_t)$  can be written as for some  $\tilde{c}$ :  $\tilde{c}(z_t, v_t)$ . For the finite delay case, the cost per-stage can further be specialized as  $\tilde{c}(x_{t-d}, v_t)$ . For the Markov case with memory, the cost function per-stage writes as  $\tilde{c}(x_{[t-m+1,t]}, v_t)$ . Now, by replacing  $\mathbb{X}$  with  $\mathbb{X}^{\max(d+1,m)}$ , let  $\pi_t \in \mathcal{P}(\mathbb{X}^{\max(d+1,m)})$  be given by

$$\pi_t(A) = P(z_t \in A | y_{[0,t]}), \quad A \in \mathcal{B}(\mathbb{X}^{\max(d+1,m)})$$

and  $\Xi_t \in \mathcal{P}(\mathcal{P}(\mathbb{X}^{\max(d+1,m)}))$  be the regular conditional measure defined by

$$\Xi_t(A) = P(\pi_t \in A | q_{[0,t-1]}), \quad A \in \mathcal{B}(\mathcal{P}(\mathbb{X}^{\max(d+1,m)})).$$

Hence, we have the following result, which is a direct extension of Theorems 2.4 and 2.5.

**Theorem 2.6:** Suppose that the partially observed source is such that, the source is Markov of order  $m$ , or there is a finite delay  $d > 0$  which is allowed at the decoder. With  $z_t = \{x_{[t-\max(d+1,m)+1,t]}\}$ ,  $y_t$  satisfies (6). Then, we have the following extensions:

- i) Any *causal composite quantization policy* can be replaced, without any loss in performance, by one which only uses  $\{\pi_t, q_{[0,t-1]}\}$  as a sufficient statistic for  $t \geq 1$ . This can be expressed as a quantization policy which only uses  $q_{[0,t-1]}$  to generate a quantizer, where the quantizer uses  $\pi_t$  to generate the quantization output at time  $t$ .
- ii) Any *optimal causal composite quantization policy* can be replaced, without any loss in performance, by one which only uses  $\{\Xi_t, \pi_t, t\}$  for  $t \geq 1$ . This can be expressed as an optimal quantization policy which only uses  $\{\Xi_t, t\}$  to generate an optimal quantizer, where the quantizer uses  $\pi_t$  to generate the quantization output at time  $t$ .  $\diamond$

For a further case where the decoder's memory is limited or imperfect, the results apply by replacing the full information at the receiver considered so far in our analysis with the limited memory under additional technical assumptions on the decoder's update of its memory (in particular, (15) in the proof of Theorem 2.5 does not apply in general). However, an equivalent result of Theorem 2.4 applies also for the limited memory setting. Such memory settings have been considered in [52], [46] and [36].

## III. MULTI-TERMINAL (DECENTRALIZED) SETUP

### A. Case with memoryless sources

Let us first consider a special, but important, case when  $\{x_t, t \geq 0\}$  is an independent and identically distributed (i.i.d.) sequence. Further, suppose that, the observations are given by

$$y_t^i = g^i(x_t, r_t^i), \quad (7)$$

for measurable functions  $g^i, i = 1, 2$ , with  $\{r_t^1, r_t^2\}$  (across time) an i.i.d. noise process. We do not require that  $r_t^1$  and  $r_t^2$  are independent for a given  $t$ . We note that our result below is also applicable when the process  $\{r_t^1, r_t^2\}$  is only independent (across time), but not necessarily identically distributed.

One difference with the general setup considered earlier in Section I is that we require the observation spaces  $\mathbb{Y}^i, i = 1, 2$ , to be finite spaces ( $\mathbb{X}$  can still be Polish).

Suppose the goal is again the minimization problem

$$\inf_{\Pi^{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi^{comp}, \gamma} \left[ \sum_{t=0}^{T-1} c(x_t, v_t) \right], \quad (8)$$

over all causal coding and receiver decision policies.

We now make a definition. In the following  $1_E$  denotes the indicator function of an event  $E$ .

**Definition 3.1:** We define the class of non-stationary memoryless team policies at  $t \geq 0$  as follows:

$$\begin{aligned} \Pi^{NSM} &:= \left\{ \Pi^{comp} : P(q_t | y_{[0,t]}) \right. \\ &= P(q_t^1 | y_t^1, t) P(q_t^2 | y_t^2, t) = 1_{\{q_t^1 = Q_t^1(y_t^1)\}} 1_{\{q_t^2 = Q_t^2(y_t^2)\}}, \\ &\left. Q_t^i : \mathbb{Y}^i \rightarrow \mathcal{M}_t^i, i = 1, 2, t \geq 0 \right\}, \quad (9) \end{aligned}$$

where, in the above,  $\{Q_t^1, Q_t^2\}$  are arbitrary measurable functions.

**Theorem 3.1:** Consider the minimization of (8). An optimal composite quantization policy over all causal policies is an element of  $\Pi^{NSM}$ . Such a policy exists.  $\diamond$

The proof is presented in the Appendix.

Hence, an optimal composite quantization policy only uses the product form admitted by a non-stationary memoryless team policy. It ignores the past observations and past quantization outputs. We note that, the proof also applies to the case when the source is memoryless, but not necessarily i.i.d.. One may ask why feedback could be useful when the source is i.i.d.. Feedback may be useful for at least two reasons: (i) Feedback can be used as a signaling mechanism for the encoders to communicate with each other (which we discuss further in the next subsections), and (ii) Feedback can provide common randomness to allow a convexification of the space of possibly randomized decentralized encoding strategies. Consider the optimization problem discussed in (8)

$$J(\Pi^{comp}) = \inf_{\gamma} E_{\nu_0}^{\Pi^{comp}, \gamma} \left[ \sum_{t=0}^{T-1} c(x_t, v_t) \right].$$

The function  $J(\Pi)$  is concave in the choice of a team policy  $\Pi$  (see Theorem 4.1 in [57] for the case with  $T = 1$ , the proof also holds for the current setting). As such, if the space of joint encoding policies is convexified by common randomness, an

optimal solution would exist at an extreme point; which in turn does not require a use of common randomness. This explains why common randomness generated by past quantization outputs does not present further benefit in the current setting.

We note before ending this subsection that if there is an entropy constraint on the quantizer outputs, then feedback might be useful for finite horizon problems as it provides common randomness, which cannot be achieved by time-sharing in a finite-horizon problem. [40] observed that randomization of two scalar quantizers (operationally achievable through time-sharing) is optimal in causal coding of an i.i.d. source subject to distortion constraints, which also applies in the side information setting of [47]. On the other hand, for the zero-delay setting, when one considers the distortion minimization problem subject to an entropy constraint, [25] observed that the distortion-entropy curve is non-convex, leading to a benefit of common-randomness for achieving points in the lower convex hull of this curve. Further relevant discussions on randomization and optimal quantizer design are present in [19] and [56].

### B. Case with Markov sources: A counterexample with signaling

We now consider Markov sources and exhibit that it is, in general, not possible to obtain a separation result of the form presented for the single-terminal case.

We will consider a two-encoder setup for the following result, where the encoders have access to the feedback from the receiver (Fig. 1). We have the following result.

**Proposition 3.1:** Consider the setup in (1)-(2) and let  $\pi_t^i = P(x_t|y_{[0,t]}^i)$ ,  $x_t \in \mathbb{X}$ ,  $i = 1, 2$ . An optimal composite quantization policy cannot, in general, be replaced by a policy which only uses  $\{\mathbf{q}_{[0,t-1]}^i, \pi_t^i\}$  to generate  $q_t^i$  for  $i = 1, 2$ .  $\diamond$

**Proof.** It suffices to produce an instance where an optimal policy cannot admit the separated structure. Toward this end, let  $z_1, z_2, z_3$  be uniformly distributed, independent, binary numbers;  $x_0, x_1$  be defined by:

$$x_0 = \begin{bmatrix} z_1 \\ z_2 \\ 0 \\ 0 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 0 \\ 0 \\ z_2 \\ z_3 \end{bmatrix},$$

such that  $x_0(1) = z_1, x_0(2) = z_2, x_0(3) = x_0(4) = 0$ . Let the observations be given as follows:

$$\begin{aligned} y_t^1 &= g^1(x_t) = x_t(1) \oplus x_t(3) \oplus x_t(4) \\ y_t^2 &= g^2(x_t) = x_t(1) \oplus x_t(2), \quad t = 0, 1. \end{aligned}$$

where  $\oplus$  is the x-or operation. That is,

$$\begin{aligned} y_0^1 &= [z_1], \quad y_0^2 = [z_1 \oplus z_2], \\ y_1^1 &= [z_2 \oplus z_3], \quad y_1^2 = [0] \end{aligned}$$

Let the cost be:

$$E \left[ (x_0(4) - E[x_0(4)|\mathbf{q}_{[0]}])^2 + (x_1(4) - E[x_1(4)|\mathbf{q}_{[0,1]}])^2 \right]$$

That is, the cost is  $E[(z_3 - E[z_3|\mathbf{q}_{[0,1]}])^2]$ , where  $q_t^i$  are the information bits sent to the decoder for  $t = 0$  and 1.

We further restrict the information rates to satisfy:  $|\mathcal{M}_0^1| = |\mathcal{M}_1^1| = |\mathcal{M}_1^2| = 2$ ,  $|\mathcal{M}_0^2| = 1$ . That is, the encoder 2 may only send information at time  $t = 1$ .

Under arbitrary causal composite quantization policies, a cost of zero can be achieved as follows: If the encoder 1 sends the value  $z_1$  to the receiver, and at time 1, encoder 1 transmits  $z_2 \oplus z_3$  and encoder 2 transmits  $z_2$  (or  $z_1 \oplus z_2$ ), the receiver can uniquely identify the value of  $z_3$ , for every realization of the random variables.

For such a source, an optimal composite policy cannot be written in the separated form, that is, an optimal policy of encoder 2 at time 1 cannot be written as  $h_1(\mathbf{q}_0, \pi_1^2)$ , for some measurable function  $h_1$ . To see this, note the following: The conditional distribution on  $x_1$  at encoder 2 at time 1 is such that the conditional measure on  $(z_2, z_3)$  is uniform and independent, that is  $P(z_2 = a, z_3 = b|z_1 \oplus z_2) = (1/4)$  for all values of  $a, b$ . If a policy of the structure of  $h_1$  is adopted, then it is not possible for encoder 2 to recall its past observation to extract the value of  $z_2$ . This is because,  $\pi_1^2$  will be a distribution only on  $z_2$  and  $z_3$ , which will be uniform and independent, given  $z_1 \oplus z_2$ . Thus, the information  $y_0^2$  will not be available in the memory and the receiver will have access to at most  $z_2 \oplus z_3$  and  $z_1$  and  $P(z_2, z_3|z_1 \oplus z_2)$  (the last variable containing no useful information). The optimal estimator will be  $E[z_3] = 1/2$ , leading to a cost of  $1/4$ .  $\diamond$

### C. Discussion: Connections with team decision theory

In this subsection, we interpret the results of the previous subsections. We first provide a brief discussion on information structures in a decentralized optimization problem: Consider a collection of decision makers (DMs) where each has access to some local information variable. Such a collection of decision makers who wish to minimize a common cost function and who has an agreement on the system (that is, the probability space on which the system is defined, and the policy and action spaces) is said to be a *team*. Such a team is *dynamic* if the information of one DM is affected by the policy of some other DM. If there is a pre-specified order of action for the DMs, the team is said to be *sequential*. Witsenhausen [50] provided the following characterization of information structures in a dynamic sequential team: Under a *centralized information structure*, all DMs have the same information. If a DM's, say  $DM^j$ , information is dependent on the policy of another DM, say  $DM^k$ , and  $DM^j$  does not have access to the information available to  $DM^k$ , this information structure is said to admit a *non-classical information structure*. A decentralized system admits a *quasi-classical information structure*, if it is not non-classical.

In a decentralized optimization problem, when the information structure is non-classical, DMs might choose to communicate via their control actions: This is known as *signaling* in decentralized control (see for example [54]).

With the characterization of information structures above, every lossy coding problem is non-classical, since a receiver cannot recover the information available at the encoder fully, while its information is clearly affected by the coding policy

of the encoder. However, in an encoding problem, the problem itself is the transmission of information. Therefore, we suggest the following: *Signaling in a coding problem is the policy of an encoder to use the quantizers/encoding functions to transmit a message to other decision makers, or to itself to be used in future stages, through the information sent to the receiver.* In the information theory literature, signaling has been employed in coding for Multiple Access Channels with feedback in [14], [12] and [44]. In these papers, the authors used active information transmission to allow for coordination between encoders.

The reason for the negative conclusion in Proposition 3.1 is that in general for an optimal policy,

$$P(q_t^i | \pi_t^i, \mathbf{q}_{[0,t-1]}, y_{[0,t-1]}^i) \neq P(q_t^i | \pi_t^i, \mathbf{q}_{[0,t-1]}), \quad (10)$$

when the encoders have engaged in signaling (in contrast with what we will have in the proof of the separation results). The encoders may benefit from using the received past observation variables explicitly.

Separation results for such dynamic team problems typically require information sharing between the encoders (decision makers), where the shared information is used to establish a sufficient statistic living in a fixed state space and which admits a controlled Markov recursion (hence, such a sufficient statistic can serve as a *state* for the decentralized system). For the proof of Theorem 2.5, we see that  $\Xi_t$  forms such a state. For the proof of Theorem 3.1, we see that information sharing is not needed for the encoders to agree on a sufficient statistic, since the source considered is memoryless. Furthermore, for the multi-terminal setting with a Markov source, a careful analysis of the proof of Theorem 3.1 (see (21), (24) and the subsequent discussion) reveals that if the encoders agree on  $P(dx_t | y_{[0,t-1]})$  through sharing their beliefs for  $t \geq 1$ , then a separation result involving this joint belief can be obtained. See [54] for a related information sharing pattern and discussions. Further results on such a *dynamic programming approach to dynamic team problems* are present in [2], [32], [39], [35], among other references.

**Remark 3.1:** In the context of multi-terminal systems, for the computation of the capacity of Multiple Access Channels with memory and partial state feedback at the encoders, a relevant discussion has been reported in [13] (section V). This is in the same spirit as our current paper in that, [13] obtains an optimality result when the channel is memoryless; and points out the difficulties arising in the case of channels with memory in view of intractability of the optimization problem: One cannot identify a finite dimensional sufficient statistic for the encoders to use. Along a relevant direction, Section III.D of [38], in the context of real-time coding, discusses the issue of the growing state space.  $\diamond$

#### IV. APPLICATION TO LINEAR QUADRATIC GAUSSIAN (LQG) ESTIMATION PROBLEMS

Consider a Linear Quadratic Gaussian setup, where a sensor quantizes its noisy information to an estimator. Let  $x_t \in \mathbb{R}^n$ ,  $y_t \in \mathbb{R}^m$ , and the evolution of the system be given by the following:

$$x_{t+1} = Ax_t + w_t,$$

$$y_t = Cx_t + r_t, \quad (11)$$

Here,  $\{w_t, r_t\}$  is a mutually independent, zero-mean Gaussian noise sequence with  $W = E[w_t w_t']$ ,  $R = E[r_t r_t']$  (where for a vector  $w$ ,  $w'$  denotes its transpose),  $x_0$  is a zero mean Gaussian variable, and  $A, C$  are matrices of appropriate dimensions. Suppose the goal is the computation of

$$\inf_{\Pi_{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi_{comp}, \gamma} \left[ \sum_{t=0}^{T-1} (x_t - v_t)' Q (x_t - v_t) \right], \quad (12)$$

with  $\nu_0$  denoting a Gaussian distribution for the initial state,  $Q > 0$  a positive definite matrix (See Figure 2).

The conditional measure  $\pi_t = P(dx_t | y_{[0,t]})$  is Gaussian for all time stages, which is characterized uniquely by its mean and covariance matrix for all time stages. We have the following.

**Theorem 4.1:** For the minimization of the cost in (12), any *causal composite quantization policy* can be replaced, without any loss in performance, by one which only uses the output of the Kalman Filter and the information available at the receiver.  $\diamond$

**Proof.** The result can be proven by considering a direct approach, rather than as an application of Theorems 2.4 and 2.5 (which require bounded costs, however, this assumption can be relaxed for this case), exploiting the specific quadratic nature of the problem. Let  $\|\cdot\|_Q$  denote the norm generated by an inner product of the form:  $\langle x, y \rangle = x' Q y$  for  $x, y \in \mathbb{R}^n$  for positive-definite  $Q > 0$ . The Projection Theorem for Hilbert Spaces implies that the random variable  $x_t - E[x_t | y_{[0,t]}]$  is orthogonal to the random variables  $\{y_{[0,t]}, q_{[0,t]}\}$ , where  $q_{[0,t]}$  is included due to the Markov chain condition

$$P(dx_t | y_{[0,t]}, q_{[0,t]}) = P(dx_t | y_{[0,t]}).$$

We thus obtain the following identity.

$$\begin{aligned} E[\|x_t - E[x_t | q_{[0,t]}]\|_Q^2] &= E[\|x_t - E[x_t | y_{[0,t]}]\|_Q^2] \\ &+ E\left[\left\|E[x_t | y_{[0,t]}] - E\left[E[x_t | y_{[0,t]}] \middle| q_{[0,t]}\right]\right\|_Q^2\right]. \end{aligned}$$

The second term is to be minimized through the choice of the quantizers. Hence, the term  $\tilde{m}_t := E[x_t | y_{[0,t]}]$ , which is computed through a Kalman Filter, is to be quantized (see Figure 2). Recall that by the Kalman Filter (see [33]) with

$$\Sigma_{0|-1} = E[x_0 x_0']$$

and for  $t \geq 0$ ,

$$\begin{aligned} \Sigma_{t+1|t} &= A \Sigma_{t|t-1} A' + W \\ &- (A \Sigma_{t|t-1} C') (C \Sigma_{t|t-1} C' + R)^{-1} (C \Sigma_{t|t-1} A'), \end{aligned}$$

the following recursion holds for  $t \geq 0$  and with  $\tilde{m}_{-1} = 0$ :

$$\begin{aligned} \tilde{m}_t &= A \tilde{m}_{t-1} \\ &+ \Sigma_{t|t-1} C' (C \Sigma_{t|t-1} C' + R)^{-1} (C A (x_{t-1} - \tilde{m}_{t-1}) + r_t). \end{aligned}$$

Thus, the pair  $(\tilde{m}_t, \Sigma_{t|t-1})$  is a Markov source, where the evolution of  $\Sigma_{t|t-1}$  is deterministic. Even though the cost to be minimized is not bounded, since  $\tilde{m}_t$  itself is a fully observed process, the proof of Theorem 2.4 can be modified to develop

the structural result that any causal encoder can be replaced with one which uses  $(\tilde{m}_t, \Sigma_{t|t-1})$  and the past quantization outputs (this result can be proven also using Theorem 1 of Witsenhausen [52], since this source is fully observed by the encoder). Likewise, the proof of Theorem 2.5 shows that, for the fully observed Markov source  $(\tilde{m}_t, \Sigma_{t|t-1})$ , any causal coder can be replaced with one which only uses the conditional probability on  $\tilde{m}_t$  and the realization  $(\tilde{m}_t, \Sigma_{t|t-1}, t)$  at time  $t$ .  $\diamond$

Thus, the optimality of Kalman Filtering allows the encoder to only use the conditional estimate and the error covariance matrix without any loss of optimality (See Figure 2), and the optimal quantization problem also has an explicit formulation. The above result is related to findings in [15] (also see [5] and [18]), and partially improves them in the direction of Markov sources.

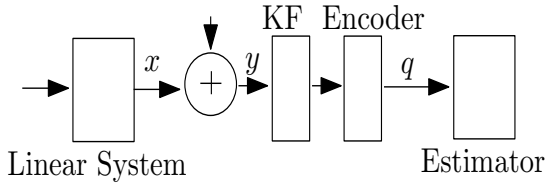


Fig. 2: **Separation of Estimation and Quantization:** When the source is Gaussian, generated by the linear system (11), the cost is quadratic, and the observation channel is Gaussian, the separated structure of the encoder above is optimal. That is, first the encoder runs a Kalman Filter (KF), and then causally encodes its estimate. For one-shot and independent observations setups, this result was observed in [5], [7], [6], [15], and [18]. Our result shows that, an extension of this result applies for the optimal causal encoding of partially observed Markov sources as well.

We note that, the above result also applies to the settings when a controller acts on the system, that is, there exists for  $u_t \in \mathbb{R}^m$  and a matrix  $B$  such that  $x_{t+1} = Ax_t + Bu_t + w_t$ . In this case, the well-known principle of separation and control in control theory allows the above results to be applicable. In particular, the conditional estimation error is not affected by the control actions.

## V. CONCLUSION

For the optimal causal coding of a partially observed Markov source, the structure of the optimal causal coders is obtained and is shown to admit a separation structure. We observed in particular that *separation of estimation (conditional probability computation) and quantization (of this probability)* applies, under such a setup. We also observed that the real-time decentralized coding of a partially observed i.i.d. source admits a separation. Such a separation result does not, in general, extend to decentralized coding of partially observed Markov sources.

The results and the general program presented in this paper apply also to coding over discrete memoryless noisy channels with noiseless feedback. We note that Walrand and Varaiya

[46] considered the noisy channel setting in their analysis in the presence of noiseless feedback.

The separation result will likely find many applications in sensor networks and networked control problems where sensors have imperfect observation of a plant to be controlled. One direction is to find explicit results on the optimal policies using computational tools. One promising approach is the expert-based systems, which are very effective once one imposes a structure on the designs, see [26] for details.

One further problem is on the existence and design of optimal quantizers. Existence of optimal quantizers, even in the context of vector quantization for  $\mathbb{R}^n$ -valued random variables, requires stringent conditions. Such proofs typically have the form of Weierstrass theorem: A lower semi-continuous function over a compact set admits a minimum. Existence results for optimal quantizers for a one-stage cost problem has been investigated by Abaya and Wise [1] and Pollard [41] for continuous cost functions which are non-decreasing in the source-reconstruction distance. [56] obtained existence results for more general cost functions under the restriction that the code bins/cells are convex and the source admits a density function. For dynamic quantizers, [58] established the existence of optimal quantization policies under the assumption that the quantizers admit the structure suggested by Theorem 2.5 for fully observed Markov sources and the codecells are convex for a class of Markov sources. Also for dynamic vector quantizers, [11] investigated the existence results when there is a bound on the quantizer bins.

Theorem 2.5 motivates the problem of optimal quantization of probability measures. This remains as an interesting problem to be investigated in a real-time coding context, with important practical consequences in control and economics applications. With a separation result paving the way for an MDP formulation, one could proceed with the analysis of [11] with the evaluation of optimal quantization policies and existence results for infinite horizon problems. Toward this direction, Graf and Luschgy, in [22] and [23], have studied the optimal quantization of probability measures.

One related question to be pursued further is the following: When is there an incentive for signaling in coding problems? When the observations are correlated for sources with memory or when the real-time coding of possibly independent sources over a general MAC type channel is considered, there may be an incentive for signaling. Further results on this will provide some light on some outstanding problems such as the capacity of MAC channels with feedback [14].

## APPENDIX

### A. Proof of Theorem 2.4

We transform the problem into a real-time coding problem involving a fully observed Markov source. At time  $t = T - 1$ , the per-stage cost function can be written as follows, where  $\gamma_t^0$  denotes a fixed receiver policy:

$$\begin{aligned} & E[c(x_t, \gamma_t^0(q_{[0,t]})) | q_{[0,t-1]}] \\ &= \sum_{\mathcal{M}_t} P(q_t = k | q_{[0,t-1]}) \left( \int_{\mathbb{X}} P(dx_t | q_{[0,t-1]}, k) \right) \end{aligned}$$



$$\begin{aligned}
& \times c(x_t, \gamma_t^0(q_{[0,t-1]}, k)) \Big) \\
&= \int_{\mathbb{X}} \sum_{\mathcal{M}_t} P(dx_t, q_t = k | q_{[0,t-1]}) c(x_t, \gamma_t^0(q_{[0,t-1]}, k)) \\
&= \int_{\mathcal{P}(\mathbb{X})} \int_{\mathbb{X}} \sum_{\mathcal{M}_t} P(dx_t, q_t = k, d\pi_t | q_{[0,t-1]}) \\
&\quad \times c(x_t, \gamma_t^0(q_{[0,t-1]}, k)) \\
&= \int_{\mathcal{P}(\mathbb{X})} \int_{\mathbb{X}} \sum_{\mathcal{M}_t} P(d\pi_t | q_{[0,t-1]}) \\
&\quad \times P(dx_t | \pi_t) P(q_t = k | \pi_t, q_{[0,t-1]}) c(x_t, \gamma_t^0(q_{[0,t-1]}, k)) \\
&= \int_{\mathcal{P}(\mathbb{X})} \sum_{\mathcal{M}_t} P(d\pi_t | q_{[0,t-1]}) P(q_t = k | \pi_t, q_{[0,t-1]}) \\
&\quad \times \int_{\mathbb{X}} P(dx_t | \pi_t) c(x_t, \gamma_t^0(q_{[0,t-1]}, k)) \\
&= E[F(\pi_t, q_{[0,t-1]}, q_t) | q_{[0,t-1]}], \tag{13}
\end{aligned}$$

where,  $\pi_t(\cdot) = P(x_t \in \cdot | y_{[0,t]})$  and

$$F(\pi_t, q_{[0,t-1]}, q_t) = \int_{\mathbb{X}} \pi_t(dx) c(x, \gamma_t^0(q_{[0,t-1]}, q_t)).$$

In the above derivation, the fourth equality follows from the property that

$$x_t \leftrightarrow P(dx_t | y_{[0,t]}) \leftrightarrow q_{[0,t]}$$

We note that  $F(\cdot, \gamma_t^0(q_{[0,t-1]}, q_t))$  is measurable on  $\mathcal{P}(\mathbb{X})$  under weak convergence topology by Theorem 2.3 and the fact that the cost is measurable and bounded.

It should be noted that for every composite quantization policy, one may define  $q_t$  as a random variable on the probability space such that the joint distribution of  $(q_t, \pi_t, q_{[0,t-1]})$  matches the characterization that  $q_t = Q_t^{comp,*}(y_{[0,t]}, q_{[0,t-1]})$ , since

$$\begin{aligned}
& P(q_t | \pi_t, q_{[0,t-1]}) \\
&= \int_{\mathbb{Y}^{t+1}} P(q_t | y_{[0,t]}, q_{[0,t-1]}) P(dy_{[0,t]} | \pi_t, q_{[0,t-1]}).
\end{aligned}$$

The final stage cost is thus written as

$$E[F(\pi_t, q_{[0,t-1]}, q_t) | q_{[0,t-1]}],$$

which is equivalent to, by the smoothing property of conditional expectation, the following:

$$E \left[ E[F(\pi_t, q_{[0,t-1]}, q_t) | \pi_t, q_{[0,t-1]}] \Big| q_{[0,t-1]} \right]$$

Now, we will apply Witsenhausen's two stage lemma [52], to show that we can obtain a lower bound for the double expectation by picking  $q_t$  as a result of a measurable function of  $\pi_t, q_{[0,t-1]}$ . Thus, we will find a composite quantization policy which only uses  $(\pi_t, q_{[0,t-1]})$  which performs as well as one which uses the entire memory available at the encoder. To make this precise, let us fix the decision function  $\gamma_t^0$  at the receiver corresponding to a given composite quantization policy at the encoder  $Q_t^{comp,*}$ , let  $t = T - 1$ , and define for every  $k \in \mathcal{M}_t$ :

$$\beta_k := \left\{ \pi_t, q_{[0,t-1]} : F(\pi_t, q_{[0,t-1]}, k) \right.$$

$$\left. \leq F(\pi_t, q_{[0,t-1]}, q'), \forall q' \neq k, q' \in \mathcal{M}_t \right\}.$$

These sets are Borel, by the measurability of  $F$  on  $\mathcal{P}(\mathbb{X})$ . Such a construction covers the domain set consisting of  $(\pi_t, q_{[0,t-1]})$  but with overlaps. It covers the elements in  $\mathcal{P}(\mathbb{X}) \times \prod_{t=0}^{T-2} \mathcal{M}_t$ , since for every element in this product set, there is a minimizing  $k \in \mathcal{M}_t$  (note that  $\mathcal{M}_t$  is finite). To avoid the overlaps, we adopt the following technique which was introduced in Witsenhausen [52]. Let there be an ordering of the elements in  $\mathcal{M}_t$  as  $1, 2, \dots, |\mathcal{M}_t|$ , and for  $k \geq 1$  in this sequence define a function  $Q_t^{comp,*}$  as:

$$\begin{aligned}
q_t &= Q_t^{comp,*}(\pi_t, q_{[0,t-1]}) = k, \\
&\text{if } (\pi_t, q_{[0,t-1]}) \in \beta_k \setminus \cup_{i=1}^{k-1} \beta_i,
\end{aligned}$$

with  $\beta_0 = \emptyset$ . Thus, for any random variable  $q_t$  appropriately defined on the probability space,

$$\begin{aligned}
& E \left[ E[F(\pi_t, q_{[0,t-1]}, q_t) | \pi_t, q_{[0,t-1]}] \Big| q_{[0,t-1]} \right] \\
&\geq E \left[ E[F(\pi_t, q_{[0,t-1]}, Q_t^{comp,*}(\pi_t, q_{[0,t-1]})) | \pi_t, q_{[0,t-1]}] \right. \\
&\quad \left. \Big| q_{[0,t-1]} \right]
\end{aligned}$$

Thus, the new composite policy performs at least as well as the original composite coding policy even though it has a restricted structure.

As such, any composite policy can be replaced with one which uses only  $\{\pi_t, q_{[0,t-1]}\}$  without any loss of performance, while keeping the receiver decision function  $\gamma_t^0$  fixed. It should now be noted that  $\{\pi_t\}$  is a Markov process: Note that the following filtering equation applies [10]

$$\begin{aligned}
& P(dx_t | dy_{[0,t]}) \\
&= \frac{\int_{x_{t-1}} P(dy_t | x_t) P(dx_t | x_{t-1}) P(dx_{t-1} | dy_{[0,t-1]})}{\int_{x_{t-1}, x_t} P(dy_t | x_t) P(dx_t | x_{t-1}) P(dx_{t-1} | dy_{[0,t-1]})},
\end{aligned}$$

and  $P(dy_t | \pi_s, s \leq t-1) = \int_{x_t} P(dy_t, dx_t | \pi_s, s \leq t-1) = P(dy_t | \pi_{t-1})$ . These imply that (see [10], see also the discussion following (15)) the following defines a Markov kernel (that is, a regular conditional probability measure):

$$P(d\pi_t | \pi_s, s \leq t-1) = P(d\pi_t | \pi_{t-1}). \tag{14}$$

We have thus obtained the structure of the encoder for the final stage. We iteratively proceed to study the other time stages. In particular, since  $\{\pi_t\}$  is Markov, we could proceed as follows (in essence using Witsenhausen's three-stage lemma [52]): For a three-stage cost problem, the cost at time  $t = 2$  can be written as, for measurable functions  $c_2, c_3$ :

$$\begin{aligned}
& E \left[ c_2(\pi_2, v_2(q_{[1,2]}), q_{[1,2]}) \right. \\
& \quad \left. + E[c_3(\pi_3, v_3(q_{[1,2]}), Q_3^{comp,*}(\pi_3, q_{[1,2]})) | q_{[1,2]}, \pi_{[1,2]}] \right. \\
& \quad \left. \Big| q_{[1,2]}, \pi_{[1,2]} \right]
\end{aligned}$$

Since

$$P(d\pi_3, q_2, q_1 | \pi_2, \pi_1, q_2, q_1) = P(d\pi_3, q_2, q_1 | \pi_2, q_2, q_1),$$

and since under  $Q_3^{comp,*}$ ,  $q_3$  is a function of  $\pi_3$  and  $q_1, q_2$ , the expectation above is equal to, for some measurable  $F_2(\cdot)$ ,  $E[F_2(\pi_2, q_2, q_1)|\pi_2, \pi_1, q_2, q_1]$ . Following similar steps as earlier, it can be established now that a composite quantization policy at time 2 uses  $\pi_2$  and  $q_1$  and is without any loss.

By a similar argument, an encoder at time  $t$ ,  $1 \leq t \leq T-1$  only uses  $(\pi_t, q_{[0,t-1]})$ . The encoder at time  $t = 0$  uses  $\pi_0$ , where  $\pi_0 = \nu_0$  is the prior distribution on the initial state.

Now that we have obtained the structure for a composite policy, we can express this as:

$$Q_t^{comp}(\pi_t, q_{[0,t-1]}) = Q^{q_{[0,t-1]}}(\pi_t), \quad \forall \pi_t, q_{[0,t-1]}$$

such that the quantizer action  $Q^{q_{[0,t-1]}} \in \mathbb{Q}(\mathcal{P}(\mathbb{X}); \mathcal{M}_t)$  is generated using only  $q_{[0,t-1]}$ , and the quantizer outcome is generated by evaluating  $Q^{q_{[0,t-1]}}(\pi_t)$  for every  $\pi_t$ .  $\diamond$

### B. Proof of Theorem 2.5

At time  $t = T-1$ , the per-stage cost function can be written as:

$$\begin{aligned} & E[c(x_t, \gamma_t(q_{[0,t]})) | q_{[0,t]}] \\ &= E\left[\int_{\mathbb{X}} P(dx_t | q_{[0,t-1]}, q_t) c(x_t, \gamma_t(q_{[0,t]}))\right] \end{aligned}$$

Thus, at time  $t = T-1$ , an optimal receiver (which is deterministic without any loss of optimality, see [9]) will use  $P(dx_t | q_{[0,t]})$  as a sufficient statistic for an optimal decision (or any receiver can be replaced with one which uses this sufficient statistic without any loss). Let us fix such a receiver policy which only uses the posterior  $P(dx_t | q_{[0,t]})$  as its sufficient statistic. Note that

$$P(dx_t | q_{[0,t]}) = \int_{\pi_t} P(dx_t | \pi_t) P(d\pi_t | q_{[0,t]})$$

and further

$$\begin{aligned} P(d\pi_t | q_{[0,t]}) &= \frac{P(q_t, d\pi_t | q_{[0,t-1]})}{\int_{\pi_t} P(q_t, d\pi_t | q_{[0,t-1]})} \\ &= \frac{P(q_t | \pi_t, q_{[0,t-1]}) P(d\pi_t | q_{[0,t-1]})}{\int_{\pi_t} P(q_t | \pi_t, q_{[0,t-1]}) P(d\pi_t | q_{[0,t-1]})} \end{aligned} \quad (15)$$

The term  $P(q_t | \pi_t, q_{[0,t-1]})$  is determined by the quantizer action  $Q_t$  (this follows from Theorem 2.4). Furthermore, given  $Q_t$ , the relation (15) is measurable on  $\mathcal{P}(\mathcal{P}(\mathbb{X}))$  (that is, in  $\Xi_t(\cdot) = P(\pi_t \in \cdot | q_{[0,t-1]})$ ) under weak convergence.

To prove this technical argument, consider the numerator in (15) and note that the function  $\kappa_B : \mathcal{P}(\mathcal{P}(\mathbb{X})) \rightarrow \mathbb{R}$  defined as  $\kappa_B(\Xi) = \Xi(B)$  is measurable under weak convergence topology as a consequence of Theorem 2.3, for each  $B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$ . By Theorem 2.1 in Dubins and Freedman [16], this implies that the relation in (15) is measurable on  $\mathcal{P}(\mathcal{P}(\mathbb{X}))$  (since the topology considered in [16] is not stronger than the weak convergence topology, the result in [16] holds in this case as well).

Let us denote the quantizer applied, given the past realizations of quantizer outputs as  $Q_t^{q_{[0,t-1]}}$ . Note that  $q_t$  is deterministically determined by  $(\pi_t, Q_t^{q_{[0,t-1]}})$  and the optimal receiver function can be expressed as  $\gamma_t^0(\Xi_t, q_t)$  (as a measurable function), given  $Q_t^{q_{[0,t-1]}}$ . The cost at time  $t = T-1$

can be expressed, given the quantizer  $Q_t^{q_{[0,t-1]}}$ , for some Borel function  $G$ , as  $G(\Xi_t, Q_t^{q_{[0,t-1]}})$ , where

$$\begin{aligned} & G(\Xi_t, Q_t^{q_{[0,t-1]}}) \\ &= \int_{\mathcal{P}(\mathbb{X})} \Xi_t(d\pi_t) \sum_{\mathcal{M}_t} 1_{\{q_t = Q_t^{q_{[0,t-1]}}(\pi_t)\}} \eta^{Q^{q_{[0,t-1]}}}(\Xi_t, q_t), \end{aligned}$$

with

$$\eta^{Q^{q_{[0,t-1]}}}(\Xi_t, q_t) = \int \pi_t(dx_t) c(x_t, \gamma_t^0(\Xi_t, q_t)) \quad (16)$$

Now, one can construct an equivalence class among the past  $q_{[0,t-1]}$  sequences which induce the same  $\Xi_t$ , and can replace the quantizers  $Q_t^{q_{[0,t-1]}}$  for each class with one which induces a lower cost among the finitely many elements in each such class, for the final time stage. Thus, an optimal quantization output may be generated using  $\Xi_t(\cdot) = P(\pi_t \in \cdot | q_{[0,t-1]})$  and  $\pi_t$ . Since there are only finitely many past sequences and finitely many  $\Xi_t$ , this leads to a Borel measurable selection of  $\pi_t$  for every  $\Xi_t$ , leading to a quantizer and a measurable selection in  $\Xi_t, \pi_t$ .

Since such a selection for  $Q_t$  only uses  $\Xi_t$ , an optimal quantization output may be generated using  $\Xi_t(\cdot) = P(\pi_t \in \cdot | q_{[0,t-1]})$  and  $\pi_t$ . Hence,  $G(\Xi_t, Q_t^{q_{[0,t-1]}})$  can be replaced with  $F_t(\Xi_t)$  for some  $F_t$ , without any performance loss.

The same argument applies for all time stages: At time  $t = T-2$ , the sufficient statistic both for the immediate cost, and the cost-to-go is  $P(dx_{t-1} | q_{[0,t-1]})$ , and thus for the cost impacting the time stage  $t = T-1$ , as a result of the optimality result for  $Q_{T-1}$ . To show that the separation result generalizes to all time stages, it suffices to prove that  $\{(\Xi_t, Q_t)\}$  is a controlled Markov chain, if the encoders use the structure above.

Toward this end, we establish that for  $t \geq 1$ , for all  $B \in \mathcal{B}(\mathcal{P}(\mathcal{P}(\mathbb{X})))$ , equation (18) (displayed on the next page) holds.

Here, (17) uses the fact that  $P(q_{t-1} | \pi_{t-1}, q_{[0,t-2]})$  is identified by  $\{P(d\pi_{t-1} | q_{[0,t-2]}), Q_{t-1}\}$ , which in turn is uniquely identified by  $q_{[0,t-2]}$  and  $Q_{t-1}$ . Furthermore, the relation in (18) defines a regular conditional probability measure since for all  $B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$ ,

$$\begin{aligned} \Xi_t(B) &= P(\pi_t \in B | q_{[0,t-1]}) \\ &= \int_{\pi_{t-1}} P(\pi_t \in B, d\pi_{t-1} | q_{[0,t-1]}) \\ &= \int_{\pi_{t-1}} P(\pi_t \in B | \pi_{t-1}) P(d\pi_{t-1} | q_{[0,t-1]}) \end{aligned}$$

is measurable in  $\Xi_{t-1}$ , given  $Q_{t-1}$  (as a consequence of the measurability of (15) in  $\Xi_t$ ). Hence, by the result of Dubins and Freedman mentioned above (Theorem 2.1 in [16]), we conclude that for any measurable function  $F_t$  of  $\Xi_t$

$$E[F_t(\Xi_t) | \Xi_{[0,t-1]}, Q_{[0,t-1]}] = E[F_t(\Xi_t), Q_t] | \Xi_{t-1}, Q_{t-1},$$

for every given  $Q_{t-1}$ . Now, once again an equivalence relationship between the finitely many past quantizer outputs, based on the equivalence of the conditional measures  $\Xi_{t-1}$  they induce, can be constructed. With the controlled Markov structure, we can follow the same argument for earlier time stages. Therefore, it suffices that the encoder uses only  $(\Xi_t, t)$

$$\begin{aligned}
& P\left(P(d\pi_t|q_{[0,t-1]}) \in B \middle| P(d\pi_s|q_{[0,s-1]}), Q_s, s \leq t-1\right) \\
&= P\left(\int_{\pi_{t-1}} P(d\pi_t, d\pi_{t-1}|q_{[0,t-1]}) \in B \middle| P(d\pi_s|q_{[0,s-1]}), Q_s, s \leq t-1\right) \\
&= P\left(\left\{\frac{\int_{\pi_{t-1}} P(d\pi_t|\pi_{t-1})P(q_{t-1}|\pi_{t-1}, q_{[0,t-2]})P(d\pi_{t-1}|q_{[0,t-2]})}{\int_{\pi_{t-1}, \pi_t} P(d\pi_t|\pi_{t-1})P(q_{t-1}|\pi_{t-1}, q_{[0,t-2]})P(d\pi_{t-1}|q_{[0,t-2]})}\right\} \in B \middle| P(d\pi_s|q_{[0,s-1]}), Q_s, s \leq t-1\right) \\
&= P\left(\left\{\frac{\int_{\pi_{t-1}} P(d\pi_t|\pi_{t-1})P(q_{t-1}|\pi_{t-1}, q_{[0,t-2]})P(d\pi_{t-1}|q_{[0,t-2]})}{\int_{\pi_{t-1}, \pi_t} P(d\pi_t|\pi_{t-1})P(q_{t-1}|\pi_{t-1}, q_{[0,t-2]})P(d\pi_{t-1}|q_{[0,t-2]})}\right\} \in B \middle| P(d\pi_{t-1}|q_{[0,t-2]}), Q_{t-1}\right) \quad (17) \\
&= P\left(\int_{\pi_{t-1}} P(d\pi_t, d\pi_{t-1}|q_{[0,t-1]}) \in B \middle| P(d\pi_{t-1}|q_{[0,t-2]}), Q_{t-1}\right) \\
&= P\left(P(d\pi_t|q_{[0,t-1]}) \in B \middle| P(d\pi_{t-1}|q_{[0,t-2]}), Q_{t-1}\right) \quad (18)
\end{aligned}$$

as its sufficient statistic for all time stages, to generate the optimal quantizer. An optimal quantizer uses  $\pi_t$  to generate the optimal quantization outputs.  $\diamond$

### C. Proof of Theorem 3.1

The proof is in three steps, (i), (ii) and (iii) below.

**Step (i)** In decentralized dynamic decision problems where the decision makers have the same objective (that is, in team problems), more information provided to the decision makers does not lead to any degradation in performance, since the decision makers can always choose to ignore the additional information. In view of this, let us relax the information structure in such a way that the decision makers now have access to all the previous observations, that is the information available at the encoders 1 and 2 are:

$$\begin{aligned}
I_t^i &= \{y_t^i, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}\} \quad t \geq 1, \quad i = 1, 2. \\
I_0^i &= \{y_0^i\}, \quad i = 1, 2.
\end{aligned}$$

The information pattern among the encoders is now the *one-step delayed observation sharing pattern*. We will show that the past information can be eliminated altogether, to prove the desired result.

**Step (ii)** The second step uses the following technical Lemma.

**Lemma A.1:** Under the relaxed information structure in step (i) above, any decentralized quantization policy at time  $t$ ,  $1 \leq t \leq T-1$ , can be replaced, without any loss in performance, with one which only uses  $(\pi_t, \mathbf{y}_t, \mathbf{q}_{[0,t-1]})$ , satisfying the following form:

$$\begin{aligned}
& P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]}) \\
&= P(q_t^1|y_t^1, \mathbf{q}_{[0,t-1]})P(q_t^2|y_t^2, \mathbf{q}_{[0,t-1]}) \\
&= 1_{\{q_t^1 = \bar{Q}^1(y_t^1, \mathbf{q}_{[0,t-1]})\}} 1_{\{q_t^2 = \bar{Q}^2(y_t^2, \mathbf{q}_{[0,t-1]})\}}, \quad (19)
\end{aligned}$$

for measurable functions  $\bar{Q}^1$  and  $\bar{Q}^2$ .  $\diamond$

**Proof.** Let us fix a composite quantization policy  $\Pi^{comp}$ . At time  $t = T-1$ , the per-stage cost function can be written as:

$$E\left[\int_{\mathbb{X}} P(dx_t|\mathbf{q}_{[0,t]})c(x_t, v_t)|\mathbf{q}_{[0,t-1]}\right] \quad (20)$$

For this problem,  $P(dx_t|\mathbf{q}_{[0,t]})$  is a sufficient statistic for an optimal receiver. Hence, at time  $t = T-1$ , an optimal receiver will use  $P(dx_t|\mathbf{q}_{[0,t]})$  as a sufficient statistic for an optimal decision as the cost function conditioned on  $\mathbf{q}_{[0,t]}$  is written as:  $\int P(dx_t|\mathbf{q}_{[0,t]})c(x_t, v_t)$ , where  $v_t$  is the decision of the receiver. Now, let us fix this decision policy at time  $t$ . We now note that (21), displayed on the next page, follows.

In (21), we use the relation  $P(dx_t|\mathbf{y}_{[0,t-1]}) = P(dx_t) =: \pi(dx_t)$ , where  $\pi(\cdot)$  denotes the marginal probability on  $x_t$  (recall that the source is memoryless). The term  $P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})$  in (21) is determined by the composite quantization policies:

$$\begin{aligned}
& P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]}) \\
&= P(q_t^1|y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})P(q_t^2|y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}) \\
&= 1_{\{q_t^1 = Q_t^{comp,1}(y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}} \\
&\quad \times 1_{\{q_t^2 = Q_t^{comp,2}(y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}}
\end{aligned}$$

The above is valid since each encoder knows the past observations of both encoders. As such,  $P(dx_t|\mathbf{q}_{[0,t]})$  can be written as, for some function  $\Upsilon$ :  $\Upsilon(\pi, \mathbf{q}_{[0,t-1]}, \mathbf{Q}_t^{comp}(\cdot))$ . Note that  $\mathbf{q}_{[0,t-1]}$  appears due to the term  $P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})$ . Now, consider the following space of joint (team) mappings at time  $t$ , denoted by  $\mathcal{G}_t$ :

$$\mathcal{G}_t = \{\Psi_t : \Psi_t = \{\Psi_t^1, \Psi_t^2\}, \Psi_t^i : \mathbb{Y}^i \rightarrow \mathcal{M}_t^i, \quad i = 1, 2\}. \quad (22)$$

For every composite quantization policy there exists a distribution  $P'$  on random variables  $(\mathbf{q}_t, \pi, \mathbf{q}_{[0,t-1]})$  such that

$$\begin{aligned}
P'(\mathbf{q}_t|\pi, \mathbf{q}_{[0,t-1]}) &= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} P(\mathbf{q}_t, \mathbf{y}_{[0,t]}|\pi, \mathbf{q}_{[0,t-1]}) \\
&= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} \left( P(q_t^1|y_{[0,t-1]}, y_t^1, \mathbf{q}_{[0,t-1]}, \pi) \right. \\
&\quad \times P(q_t^2|y_{[0,t-1]}, y_t^2, \mathbf{q}_{[0,t-1]}, \pi) \\
&\quad \times P(y_t^1, y_t^2)P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]}) \left. \right) \quad (23)
\end{aligned}$$

$$\begin{aligned}
P(dx_t|\mathbf{q}_{[0,t]}) &= \sum_{\mathbb{Y}^{t+1}} P(dx_t, \mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) = \frac{\sum_{\mathbb{Y}^{t+1}} P(dx_t, \mathbf{q}_t, \mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t-1]})}{P(\mathbf{q}_t|\mathbf{q}_{[0,t-1]})} \\
&= \frac{\sum_{\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)P(dx_t|\mathbf{y}_{[0,t-1]})P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}{\int_{\mathbb{X}, \mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)P(dx_t|\mathbf{y}_{[0,t-1]})P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})} \\
&= \frac{\sum_{\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)\pi(dx_t)P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}{\int_{\mathbb{X}, \mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)\pi(dx_t)P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})} \quad (21)
\end{aligned}$$

Furthermore, with every composite quantization policy and every realization of  $\mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}$ , we can associate an element in the space  $\mathcal{G}_t$ ,  $\Psi_{\mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}}$ , such that the induced stochastic relationship in (23) can be obtained:

$$\begin{aligned}
P'(\mathbf{q}_t|\pi, \mathbf{q}_{[0,t-1]}) &= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} P(\mathbf{q}_t, \mathbf{y}_{[0,t]}|\pi, \mathbf{q}_{[0,t-1]}) \\
&= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} \mathbf{1}_{\{\Psi_{\mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}}(y_t^1, y_t^2) = (q_t^1, q_t^2)\}} \\
&\quad \times P(y_t^1, y_t^2)P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]})
\end{aligned}$$

We can thus, express the cost, for some measurable function  $F$  in the following way:

$$E[F(\pi, \mathbf{q}_{[0,t-1]}, \Psi)|\pi, \mathbf{q}_{[0,t-1]}],$$

where

$$\begin{aligned}
P(\Psi|\pi, \mathbf{q}_{[0,t-1]}) &= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^t} \mathbf{1}_{\{\Psi = \Psi_{\mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}}\}} P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]})
\end{aligned}$$

Now let  $t = T - 1$  and define for every realization  $\Psi_t = (\Psi_t^1, \Psi_t^2) \in \mathcal{G}_t$  (with the decision policy considered earlier fixed):

$$\begin{aligned}
\beta_{\Psi_t} &:= \left\{ \pi, \mathbf{q}_{[0,t-1]} : \right. \\
&\quad \left. F(\pi, \mathbf{q}_{[0,t-1]}, \Psi_t) \leq F(\pi, \mathbf{q}_{[0,t-1]}, \Psi_t') \right. \\
&\quad \left. \forall ((\Psi_t^1)', (\Psi_t^2)') \in \mathcal{G}_t \right\}.
\end{aligned}$$

As we had observed in the proof of Theorem 2.4, such a construction covers the domain set consisting of  $(\pi, \mathbf{q}_{[0,t-1]})$  but possibly with overlaps. Note that for every  $(\pi, \mathbf{q}_{[0,t-1]})$ , there exists a minimizing function in  $\mathcal{G}_t$ , since  $\mathcal{G}_t$  is a finite set. In this sequence, let there be an ordering of the finitely many elements in  $\mathcal{G}_t$  as  $\{\Psi_t(1), \Psi_t(2), \dots, \Psi_t(k), \dots\}$ , and define a function  $\mathbf{T}_t^*$  as:

$$\begin{aligned}
\Psi_t(k) &= \mathbf{T}_t^*(\pi, \mathbf{q}_{[0,t-1]}), \\
\text{if } \left( \pi, \mathbf{q}_{[0,t-1]} \right) &\in \beta_{\Psi_t(k)} - \cup_{i=0}^{k-1} \beta_{\Psi_t(i)},
\end{aligned}$$

with  $\beta_{\Psi_t(0)} = \emptyset$ . Thus, we have constructed a policy which performs at least as well as the original composite quantization policy. It has a restricted structure in that it only uses  $(\pi, \mathbf{q}_{[0,t-1]})$  to generate the team action and the local information  $y_t^1, y_t^2$  to generate the quantizer outputs.

Now that we have obtained the structure of the optimal encoders for the last stage, we can sequentially proceed to

study the other time stages. Note that given a fixed  $\pi, \{(\pi, \mathbf{y}_t)\}$  is i.i.d. and hence Markov. Now, define  $\pi'_t = (\pi, \mathbf{y}_t)$ . For a three-stage cost problem, the cost at time  $t = 2$  can be written as, for measurable functions  $c_2, c_3$ :

$$\begin{aligned}
&c_2(\pi'_2, v_2(\mathbf{q}_{[1,2]})) \\
&+ E[c_3(\pi'_3, v_3(\mathbf{q}_{[1,2]}), Q_3(\pi'_3, \mathbf{q}_{[1,2]}))|\pi'_2, \mathbf{q}_{[1,2]}]
\end{aligned}$$

Since  $P(d\pi'_3, \mathbf{q}_{[1,2]}|\pi'_2, \pi'_1, \mathbf{q}_{[1,2]}) = P(d\pi'_3, \mathbf{q}_{[1,2]}|\pi'_2, \mathbf{q}_{[1,2]})$ , the expression above is equal for some  $F_2(\pi'_2, \mathbf{q}_2, \mathbf{q}_1)$  for some measurable  $F_2$ . By a similar argument, an optimal composite quantizer at time  $t, 1 \leq t \leq T - 1$  only uses  $(\pi, \mathbf{y}_t, \mathbf{q}_{[0,t-1]})$ . An optimal (team) policy generates the quantizers  $Q_t^1, Q_t^2$  using  $\mathbf{q}_{[0,t-1]}, \pi$ , and the quantizers use  $\{y_t^i\}$  to generate the quantizer outputs at time  $t$  for  $i = 1, 2$ .  $\diamond$

**Step (iii)** The final step will complete the proof. At time  $t = T - 1$ , an optimal receiver will use  $P(dx_t|\mathbf{q}_{[0,t]})$  as a sufficient statistic for the optimal decision. We now observe that

$$\begin{aligned}
P(dx_t|\mathbf{q}_{[0,t]}) &= \sum_{\mathbb{Y}^{t+1}} P(dx_t|\mathbf{y}_{[0,t]})P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) \quad (24) \\
&= \sum_{\mathbb{Y}^{t+1}} P(dx_t|\mathbf{y}_t)P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) \\
&= \sum_{\mathbb{Y}} P(dx_t|\mathbf{y}_t) \sum_{\mathbb{Y}^t} P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) \\
&= \sum_{\mathbb{Y}} P(dx_t|\mathbf{y}_t)P(\mathbf{y}_t|\mathbf{q}_{[0,t]})
\end{aligned}$$

Thus,  $P(dx_t|\mathbf{q}_{[0,t]})$ , is a function of  $P(\mathbf{y}_t|\mathbf{q}_{[0,t]})$ . Now, let us note that

$$\begin{aligned}
P(\mathbf{y}_t|\mathbf{q}_{[0,t]}) &= \frac{P(\mathbf{q}_t, \mathbf{y}_t|\mathbf{q}_{[0,t-1]})}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t, \mathbf{y}_t|\mathbf{q}_{[0,t-1]})} \\
&= \frac{P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|\mathbf{q}_{[0,t-1]})}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|\mathbf{q}_{[0,t-1]})} \\
&= \frac{P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t)}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t)}, \quad (25)
\end{aligned}$$

where the term  $P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})$  is determined by the quantizer team action  $\mathbf{Q}_t^{comp}$ . As such, the cost at time  $t = T - 1$  can be expressed as a measurable function  $G(P(\mathbf{y}_t), \mathbf{Q}_t)$ . Thus, it follows that, an optimal quantizer policy at the last stage,  $t = T - 1$  may only use  $P(\mathbf{y}_t)$  to generate the quantizers, where the quantizers use the local information  $y_t^i$  to generate the quantization output. At time  $t = T - 2$ , the sufficient statistic for the cost function is  $P(dx_{t-1}|\mathbf{q}_{[0,t-1]})$  both for the immediate cost, and the cost-to-go, that is the cost impacting the time stage  $t = T - 1$ , as a result of the optimality result

for  $Q_{T-1}$  and the memoryless nature of the source dynamics. The same argument applies for all time stages.

Hence, any policy without loss can be replaced with one in  $\Pi^{NSM}$  defined in (9). Since there are finitely many policies in this class, an optimal policy exists.  $\diamond$

**Acknowledgements** Discussions with Professors Ramon van Handel, Tamás Linder and Naci Saldi on the contents of the paper are gratefully acknowledged. The detailed reviews of two anonymous reviewers and the suggestions of the associate editor have led to significant improvement in the presentation.

## REFERENCES

- [1] E. A. Abaya and G. L. Wise, "Convergence of vector quantizers with applications to optimal quantization, *SIAM Journal on Applied Mathematics*, vol. 44, pp. 183–189, 1984.
- [2] M. Aicardi, F. Davoli and R. Minciardi, "Decentralized optimal control of Markov chains with a common past information set", *IEEE Transactions on Automatic Control*, vol. 32, pp. 1028–1031, Nov. 1987.
- [3] A. Anastasopoulos, "A sequential transmission scheme for the multiple access channel with noiseless feedback," *Proc. Annual Allerton Conference*, September 2009, IL, USA.
- [4] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide, third edition*, Springer-Verlag, Berlin, 2006.
- [5] E. Ayanoglu and R. M. Gray, "The design of joint source and channel trellis waveform coders," *IEEE Transactions on Information Theory*, vol. 33, no. 6, pp. 855–865, Nov. 1987.
- [6] R. Bansal and T. Başar, "Solutions to a class of linear-quadratic-Gaussian LQG stochastic team problems with nonclassical information", *Systems and Control Letters*, vol. 9, pp. 125 – 130, 1987
- [7] T. Berger, *Rate Distortion Theory*, Englewood Cliffs, NJ: Prentice- Hall, 1971.
- [8] V. I. Bogachev, *Measure Theory*, Springer-Verlag, Berlin, 2007.
- [9] D. Blackwell, "Memoryless strategies in finite-stage dynamic programming," *Annals of Mathematical Statistics*, vol. 35, no. 2, pp. 863–865, 1964.
- [10] V. S. Borkar and A. Budhiraja, "A further remark on dynamic programming for partially observed Markov processes", *Stochastic Processes and their Applications*, vol. 112, pp. 79–93, 2004.
- [11] V. S. Borkar, S. K. Mitter, and S. Tatikonda, "Optimal sequential vector quantization of Markov sources," *SIAM J. Control and Optimization*, vol. 40, pp. 135–148, 2001.
- [12] S. Bross and A. Lapidoth, "An improved achievable region for the discrete memoryless two-user multipleaccess channel with noiseless feedback", *IEEE Trans Inf Theory*, vol. 51, pp. 811–833, March 2005.
- [13] G. Como and S. Yüksel, "On the capacity of memoryless finite state multiple access channels with asymmetric state information at the encoders, *IEEE Transactions on Information Theory*, vol. 57, pp. 1267 – 1273, March 2011
- [14] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Transactions on Information Theory*, vol. 27, pp. 292–298, May 1981
- [15] R. L. Dobrushin and B. S. Tsybakov, "Information transmission with additional noise, *IRE Transactions on Information Theory*, vol. 18, pp. 293–304, 1962.
- [16] L. Dubins and D. Freedman, "Measurable sets of measures", *Pacific J. Math.*, vol. 14, pp. 1211–22, 1964.
- [17] H. Ebeid and T. P. Coleman, "Source coding with feedforward using the posterior matching scheme", in *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, Austin, Texas, U.S.A., June 2010.
- [18] T. Fine, "Optimum mean-square quantization of a noisy input", *IEEE Transactions on Information Theory*, vol. 11, pp. 293–294, Apr. 1965.
- [19] G. Gabor, G. Szekeres and Z. Györfi, "On the Gaarder-Slepian "Tracking System" conjecture, *IEEE Transactions on Information Theory*, vol. 37, pp. 1165–1168, July 1991.
- [20] J. K. Ghosh, and R. V. Ramamoorthi, *Bayesian Nonparametrics*, Springer, New York, 2003.
- [21] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Transactions on Information Theory*, vol. 33, pp. 773–787, Nov. 1987.
- [22] S. Graf and H. Luschgy, "Quantization for probability measures in the Prohorov metric", *Theory Probab. Appl.*, vol. 53, pp. 216–241, 2009.
- [23] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, Springer, 2000.
- [24] R. M. Gray, S. Boyd, and T. Lookabaugh, "Low rate distributed quantization of noisy observations, *Proc. Allerton Conf. Communication, Control and Computing*, Urbana-Champaign, IL, pp. 354–358, 1985.
- [25] A. Györfy and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source", *IEEE Transactions on Information Theory*, vol. 46, pp. 2704–2711, Nov. 2000.
- [26] A. Györfy, T. Linder and G. Lugosi, "Tracking the best quantizer," *IEEE Transactions on Information Theory*, vol. 54, pp. 1604–1625, April. 2008.
- [27] Y. C. Ho, "Team decision theory and information structures, *Proc. IEEE*, vol. 68, pp. 644–654, 1980.
- [28] M. Huang and S. Dey, "Dynamic quantizer design for hidden Markov state estimation via multiple sensors with fusion center feedback", *IEEE Tran. Signal Processing* 54: 2887–2896, 2006.
- [29] T. Linder and G. Lugosi, "A zero-delay sequential scheme for lossy coding of individual sequences," *IEEE Transactions on Information Theory*, vol. 47, pp. 2533–2538, Sep. 2001.
- [30] T. Linder and R. Zamir, "Causal coding of stationary sources and individual sequences with high resolution," *IEEE Transactions on Information Theory*, vol. 52, pp. 662–680, Feb. 2006.
- [31] Y. Kaspi and N. Merhav, "Structure theorem for real-time variable-rate lossy source encoders and memory-limited decoders with side information", in *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, Austin, Texas, U.S.A., June 2010.
- [32] B. Kurtaran, "Corrections and extensions to 'Decentralized control with delayed sharing information pattern *IEEE Transactions on Automatic Control*, vol. 24, pp. 656–657, Aug. 1979.
- [33] H.J. Kushner, *Introduction to Stochastic Control Theory*. Holt, Rinehart and Winston, New York, 1972.
- [34] A. Mahajan, *Sequential decomposition of sequential teams: applications to real-time communication and networked control systems*, Ph.D. Dissertation, University of Michigan, Ann Arbor, September 2008.
- [35] A. Mahajan, "Optimal decentralized control of coupled subsystems with control sharing", *Proc. 50th IEEE Conference on Decision and Control*, Dec 12–15, 2011.
- [36] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication with noisy feedback", *IEEE Journal on Special Areas in Communications*, vol. 28, pp. 580–595, May 2008.
- [37] A. Mahajan and D. Teneketzis, "Optimal design of sequential real-time communication systems", *IEEE Transactions on Information Theory*, vol. 55, pp. 5317–5338, November 2009.
- [38] A. Nayyar and D. Teneketzis, "On the structure of real-time encoding and decoding functions in a multi-terminal communication system", *IEEE Transactions on Information Theory*, vol. 57, pp. 6196–6214, September 2011.
- [39] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures", *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1606–1620, July 2011.
- [40] D. L. Neuhoff and R. K. Gilbert, "Causal source codes", *IEEE Transactions on Information Theory*, 28:701–713, September 1982.
- [41] D. Pollard, "Quantization and the method of  $k$ -means, *IEEE Transactions on Information Theory*, vol. 28, pp. 199–205, 1982.
- [42] S. Tatikonda, *Control Under Communications Constraints*, Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2000.
- [43] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication, *IEEE Transactions on Information Theory*, vol. 52, pp. 4017–4035, September 2006.
- [44] R. Venkataramanan and S. S. Pradhan, "A new achievable rate region for the discrete memoryless multiple-access channel with feedback", *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, Seoul, South Korea, June 2009.
- [45] R. Venkataramanan and S. S. Pradhan, "Source coding with feed-forward: Rate-distortion theorems and error exponents for a general source", *IEEE Transactions on Information Theory*, vol. 53, pp. 2154–2179, June 2007.
- [46] J. C. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Transactions on Information Theory*, 19:814–820, November 1983.
- [47] T. Weissman and N. Merhav, "On causal source codes with side information, *IEEE Transactions on Information Theory*, vol. 51, pp. 4003–4013, November 2005.
- [48] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels, *IEEE Transactions on Communications*, vol. 57, pp. 1693–1700, June 2009.

- [49] L. Stettner, "Ergodic control of partially observed Markov control processes with equivalent transition probabilities", *Applicaciones Mathematicae*, vol. 22, pp. 2538, 1993.
- [50] H. S. Witsenhausen, "The intrinsic model for discrete stochastic control: Some open problems," in *Control Theory, Numerical Methods and Computer System Modelling*, ser. Lecture Notes in Economics and Mathematical Systems, A. Bensoussan and J. L. Lions, Eds. Springer Verlag, 1975, vol. 107, pp. 322–335.
- [51] H. S. Witsenhausen, "Indirect rate-distortion problems", *IEEE Transactions on Information Theory*, vol. 26, pp. 518-521, Sept. 1980.
- [52] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell Syst. Tech. J.*, 58:1437-1451, July/August 1979.
- [53] H. S. Witsenhausen, "Equivalent stochastic control problems, *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 3-11, Springer-Verlag, 1988.
- [54] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern", *IEEE Transactions on Automatic Control*, 55:2773-2786, Dec. 2009.
- [55] S. Yüksel, T. Başar and S. P. Meyn "Optimal causal quantization of Markov Sources with distortion constraints", *Proc. Information Theory and Applications Workshop (ITA)*, San Diego, CA, U.S.A., Jan. 2008.
- [56] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control", *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 864-887, 2012.
- [57] S. Yüksel and T. Linder, "Optimization and convergence of observation channels and quantizers in stochastic control", *Proc. Annual Allerton Conference on Communications, Control and Computing*, pp. 286–293, Monticello, IL, 2010.
- [58] S. Yüksel and T. Linder, "On Optimal Zero-Delay Quantization of Vector Markov Sources", *Proc. IEEE Conference on Decision and Control*, 2012.

**Serdar Yüksel** received his BSc degree in Electrical and Electronics Engineering from Bilkent University in 2001; MS and PhD degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively. He was a post-doctoral researcher at Yale University for a year before joining Queen's University as an assistant professor of Mathematics and Engineering at the Department of Mathematics and Statistics. His research interests are on stochastic and decentralized control, information theory and applied probability.